Programed Instruction

October, 1961, Volume 1
Bulletin Number 2

Summer; and the Field This Fall

As the new academic year began, at least seven American colleges and universities were offering courses dealing with programed instruction; the International Cooperation Administration and the American Management Association were investigating their use for very different purposes; school systems across the country were experimenting with programs and the Center was catching its breath after a summer packed with conferences, public appearances, and sold-out workshops. [It was a hot summer—for programing, too.]

Some 208 educators attended CPI workshops in Chicago, New York, Newton, Providence, Tampa, and Rochester. They included elementary, secondary, and college teachers from virtually all academic disciplines. Department heads, reading specialists, principals, superintendents, and training directors for government and industry were among the "students." Taking the group as a whole, the average number of years of teaching experience was 12.5.

At the 69th annual convention of the American Psychological Association, held in New York over the long Labor Day week end, seven symposia discussed aspects of programed instruction and seven other meetings heard readings of papers.

Many sessions, although not specifically listing the topic related directly to the field, not the least of which were given by the Division of Experimental Psychology. Prior to the convention, a memorandum was circulated, requesting the establishment of a new APA division to deal with educational technology, including "teaching machines" and programed instruction. Following an open meeting, the convention decided to refer the matter to committee.

One of the summer's displays of programing progress came in August at a Conference - Exhibit on Programed Learning and Teaching Machines, sponsored by the American Management Association in New York. On display were devices from the simplest of pressed plastic forms to the most complex of electronic devices, all labeled as teaching machines. The programs offered on these machines, *inter alia,* taught salesmanship, operation of punch-card devices, the wiring of electronic circuits, and the means of solving Chinese puzzles. At one morning session, debaters argued the relative merits of "The Linear Programing Approach" and "The Branching Programing Approach," with a follow-up discussion on these and other programing methods. A similar conference will be held in Los Angeles in November.

In Washington, the American Council on Education and the I.C.A. jointly sponsored a four-day conference on "New Learning Processes for Developing Nations." The 87 participants included leading figures from government, education, industry, and the military. Four of the six policy committees at the conference recommended exploration of programed instruction as a way to hasten the educational progress of underdeveloped countries.

Relieved of daily classroom duties, a number of teachers and professors all over the country worked to program the subject-matter of their own specialties. The National Science Foundation awarded a grant to the School Mathematics Study Group to help the group program its imaginative approach to elementary and secondary math teaching; and SMSG held a special meeting in New Haven to help its members familiarize themselves with the theory and practice of programing. At Princeton Theological Seminary, a program in Biblical Hebrew was completed for the Didak 501 teaching machine, under the slogan, "The teaching of ancient languages goes modern!"

Research reporting the actual classroom use of programed instruction became available in greater quantity.

A Denver study of 250 tenth-

Continued on page 2

Inside Opinion

QUESTION: Can a program's quality be measured by how well the students do on existing standardized achievement tests?

ANSWER: Before one uses any standardized achievement test for purposes of evaluation, it is essential that the content of the test be carefully compared with the content of the course in conjunction with which it is to be used. To the extent that what is measured by an existing standardized test of high technical excellence corresponds to the content and objectives of the course, such a test would indeed provide a useful tool in an evaluation study. This statement would be true irrespective of the **medium** through which the course material is presented, whether it be television, self-instructional devices, or teacher-in-class.

Frederick L. Ferris, Jr.

Director of Curriculum Studies Educational Testing Service, Member of ETS Advisory Committee for the Study of Programed Instruction.

ANSWER: Existing standardized achievement tests may provide one kind of criterion for the evaluation of programs. But few, if any, of the standardized tests are completely suitable as criterion measures. All too often, large-scale achievement tests fail to distinguish between the effects of instruction and the effects of general ability and motivation. Most standardized achievement tests are constructed with an eye to selection procedures rather than as measures of instructional effects. As a consequence, the test items may be confounded and thus not sensitive to specific instructional effects.

Furthermore, the standardized achievement tests may present a danger as criteria in that they can become a standard for education, sometimes fixed in terms of conventions rather than in terms of absolute standards of excellence. When this happens, the test becomes the criterion and the concepts of educational mastery are sacrificed to a static testing procedure. In view of the new developments in curriculum and in view of the subject-matter-mastery emphasis of programed instruction, the standardized achievement tests should not be considered as adequate criterion measures by themselves.

Robert E. Silverman

Chairman, Psychology Department, University College of Arts and Science, New York University. Author and Researcher of programed instructional materials.

Continued on page 2

The Center for Programed Instruction, Inc. hereby grants to all non-profit organizations, including schools with a regular curriculum, the right to reproduce any part of this work. Reproduction by any other person, individual, corporate or otherwise, of any part of this work, without the permission of the Center is stictly prohibited.

THIS FALL (Cont. from page 1)

grade pupils reported that the *English 2600* program had been particularly effective, by both standardized and non standardized measures, with students in a high-ability group. Lakewood, Ohio, reported success with the use of the Temac Algebra I materials in a summer "algebra-improvement" class. One teacher said that "he came to know the students in his class better as individuals than in any summer school class he had taught before."

In Ulster County, New York, 280 students ranging from 7th to 11th grades took tests to determine their retention of material taught in a modern mathematics program. No significant drop in test scores was observed at 2, 4, 6, 8, or 10 week intervals, but a small, significant drop was noted at the 12th week. Further tests will be given at the six-month mark.

Colleges offering courses in programed instruction, as of last report, include Oregon College of Education, Cornell (Industrial Training), Teachers College, Columbia, Rochester, Syracuse, Southern California (Seminar in Training), and Pittsburgh, where a fall workshop will be run by the American Institute for Research. Numerous PTA and other groups will hold local study sessions, and conferences. In mid-November, Purdue University will host a three-day conference on programed instruction.

Some 40,000 subscriptions to this bulletin were entered by the not quite 150,000 to whom the first issue was sent. We are highly appreciative and humbly thankful for the affirmative feedback. Engaged in school work? Not now receiving a copy of this bulletin? Drop a post card with your name and address to the Center.

Programed Instruction is published by The Center for Programed Instruction, Inc., 365 West End Avenue, New York 24, N. Y.

Oct. 1961 Volume 1

Bulletin No. 2

Robert T. Filop
Editor

The Cost of publishing this Bulletin was paid in part under a grant from the Carnegie Corporation of New York.

Printed in U.S.A.

INSIDE OPINION (Cont. from page 1)

QUESTION: Does the introduction of programed instruction necessitate a major change in classroom administration?

ANSWER: With respect to organizing classroom routine, providing equipment and physical facilities, and planning the daily schedule, programed instruction will bring no marked changes in classrooms where individualized teaching is already part of the pattern.

Excellent precedents suggest ways of handling the new medium. Individualized teaching of reading now takes place in many classrooms. Art work is organized on an individual basis. Pupils work singly or in small groups on special projects. In small high schools, two classes in different subjects may be studying in the same room under the guidance of one teacher. A major concern will be to help the learner relate the outcomes of any programed instruction to overall educational goals.

With regard to facilities, I predict that except for special situations, the classroom will not consist of cubicles and machines. There may be several teaching machines, one or two small projectors, and a variety of programs supplementing familiar instructional materials.

Max R. Brunstetter

Professor of Education, Teachers College, Columbia University, Coordinator of Summer Workshop in Programed Instruction.

ANSWER: While programed instruction does not necessitate major change within the classroom it does make possible — even invite — such change. The potentials are great and highly intriguing. The competent, resourceful teacher, who has resigned in despair at the attempt to utilize the best techniques of individualized instruction, should find immense possibilities through the use of programed materials. For example, the elementary algebra class can be released from its age-old lock-step of everyone, all together, ready-or-not, moving from point to point. Through the use of programed instruction materials, the secondary English teacher can reduce the endless, largely unproductive hours expended on formal grammar. Programed materials in such areas as map reading and basic geographical concepts might, once more, release precious class hours. Resourceful teachers would be invited to use the time and the knowledge gained in new ways.

There is nothing inherent in the nature of programed instruction which forces teachers to new devices or new classroom organizational patterns. Those without a dream, with myopic vision, without resources will readily reduce it to the status of excellent workbook material. There are, however, teachers who will find it urging them in the direction of an almost endless variety of new ways of organizing instruction within their classrooms.

Henry O. Brunner

Principal, Memorial Place Junior High School, Manhasset, New York, where programmed instructional materials were used, Spring, 1961.

Responses

The following were made by participants in the seven summer workshops conducted by the Center:

"I have been amazed by the variety and novelty (interesting) in programs when done by teachers. Their programs with proper editing seem to be better than the commercial ones."

Mathematics Teacher, New Jersey

"The workshop has been the greatest stimulation in education that I have had in the past ten years. Perhaps I should not say the workshop *per se* but the concept of programing and its implications for education."

English Professor, Oregon

"The *place* of a program in the program in the classroom has changed for me. But I didn't think of it as a *tool* for teaching but am now quite convinced that this is one of its most valuable assets."

High School English Teacher, Oregon

"I was sorry we were left to discover for ourselves that a certain degree of creativity can be incorporated into good programing. For at least a week I felt that the limitations discouraged any teaching beyond rote learning; that there was only *one way* to program."

Social Science Teacher, Oregon

"We would have enjoyed a frank discussion regarding a shopper's view on the idea of programing — the most valuable experiences were actually in constructing programs."

Elementary School Teacher, Virginia

CPI TO GATHER, U. S. O. OF E. TO PUBLISH DATA ON PROGRAM USE

The Center began work late in September on an ANNOTATED COMPILATION OF PROGRAMED INSTRUCTION USAGE, to be published under the auspices of the U.S. Office of Education. Outlines of research will be reported in a form currently being developed by a group of educators who have employed and tested programs in their school systems. This cooperative panel consists of: Dr. Henry Brickell, Manhasset, N.Y., Dr. J. Bernard Everett, Newton, Mass., Dr. George Fernandez, Ulster County, N.Y., Dr. Joseph O. Loretan and Dr. Shelly Umans, New York City Schools, Dr. Lloid Jones, Denver, Dr. Ernest Weinrich, Huntington, N.Y., Dr. Reginald Neuwein, Cleveland. The "Patterns of Use" questionnaire will be in the mail to the Bulletin mailing list shortly. Anyone wishing to report completed or current research studies using programed instructional materials is invited to request blanks from CPI. An annual supplement to this compilation will be published as part of the Journal of Programed Instruction announced elsewhere in this issue.

PROGRAMED INSTRUCTION - EXPANDABLE OR EXPENDABLE?

Innovations, educational or otherwise, are apt to provoke controversy. Therefore, it should surprise no one that a classroom innovation as radical as programed instruction and its handmaiden the teaching machine has prompted a good deal of controversy during the past year. Arguments about programed instruction might be grouped into two separate categories.

Into the first category we may put questions regarding the ultimate value and place of programed instruction in education. Not only such questions as, "Is programed instruction better than what we now have?" but, "Even if it can produce immediate results will its long range effects be desirable?" "Furthermore, how should it be used?" "Where does programed *instruction* fit in a total *educational program?*" As fascinating, valid, and important as such questions may be, they are clearly going to be answered, not through debate, but by judicious use of and careful experimentation with *programs in schools.*

The second category includes such questions as the relative merits of the methods of programing now in vogue. Here we find long and inevitably fruitless arguments over such questions as, "Are multiple choice questions better than constructed response programs?" "Should a programer adjust to individual differences through the use of *branches?*" "Should he program from the general to the particular or *vice versa?*" This last question has prompted the invention of the name "ruleg" (rule and then examples) and "egrule" (examples and then rule) to describe the opposite approaches would be nothing more than amusing if the two camps didn't take themselves quite so seriously.

"Domestic Squabbles"

If the questions in the first category are valid, important, and generated by the educator's desire and need to know about programed instruction, many of the questions in the second category are at least confusing, if not invalid and detrimental to the growth of the field.

Neither the number nor the intensity of these "domestic squabbles" is likely to decrease in the immediate future. It may be valuable to examine some of the characteristics, causes, and possible effects of this "internal" controversy on the future of programed instruction.

Controversy carried on with accurate information and motivated by a desire to clarify issues, rather than to compete for prestige or for a market, is essential to the scientific process. Since programed instruction is undeniably an outgrowth of basic research in a behavioral science, one would hope to see a certain amount of the rigorous experimental attitude of the scientist maintained by those who have taken upon themselves the responsibility of translating the results of classroom. But a good deal of this rigorous experimental attitude has been "lost in translation," and many of the "translators" seem to have forgotten that they once spoke a common language. Nor has their "Tower of Babel" been any lofty cooperative enterprise; rather, it has been the job of competing for a major market.

Apparently life in the market place requires this kind of bravado. As John Dewey warned in 1929: "To such persons science is of value because it puts a stamp of approval upon this or that specific procedure. It is very easy for science to be regarded as a guarantee that goes with the sale of goods rather than as a light to the eyes and a lamp to the feet." Or, as James Holland of Harvard has recently stated in *Contemporary Psychology*: "We (the experimental scientists) have promised that research would enable an analysis of all prerequisite component skills and that we would be able to arrive at a sequence of carefully designed and arranged tasks that would develop these skills; now we seem to be translating these skills into procedures for expedient programing . . . . publishers promise money and fame for converting textbooks into programs. Only wisecracks about ping-pong-playing pigeons are in store for the scientist who continues to ignore the teaching machine's true destiny."

No one, including Dr. Holland, can say what the "true destiny" of the teaching machine and programed instruction will be. No matter how bright a promise the future holds, however, it is only going to be dimmed by the parochialism of any vested interest, commercial or theoretical, which asserts that a particular type of programing is best. Such an assertion implies that all things can be taught in the same way to all people.

The Blindmen and The Elephant . . . .

Those who would describe the whole of programed instruction solely in terms of their own methodology of programing bring to mind the familiar story of the blind men and the elephant. Whereas the blind men had no means of combining each other's experience into a realistic description of the elephant, however, programers need only open their eyes to see the relationship between the men and the elephant. The blind men had no way to combine each other's experience into a realistic description of the elephant, programers need only open their eyes to see the relationship between the part of programed instruction they are familiar with and the whole they must begin to make room for. Of course no one yet knows precisely what the relationship will be between the various types of programing that now exist but those who consider themselves professionals in this field should not allow vested interests or personal bias to dictate what this relationship should be. We must agree with Dr. Holland and demand that the promise of the behavioral scientist be fulfilled. The true professionals in the programing field must devote themselves to research to help all programers know what type of programing is best suited for teaching particular skills to particular students. Then and only then will the field of programed instruction be able to fulfill its potential. To expect one type of programing to teach all things equally well is as unrealistic as expecting one teacher to be able to teach all subjects with equal facility. But research in programed instruction to date has too often been asked to find out what is better *than* what, rather than what is better *for* what purposes. As Dr. Holland puts it: "We find comparisons of programed instruction with conventional methods which lead their authors to overgenerali-

Continued on page 6

CHOOSING A PROGRAM:

How to Start

(Ed. Note: The Educational Testing Service, under a grant from the Carnegie Corporation, is planning a study to identify valid criteria for the evaluation of programs. A joint committee of the A.E.R.A., A.P.A., and D.A.V.I. is also exploring this subject, as reported in the last issue of the Bulletin. This article reflects the Center's recommendations pending the establishment of validated criteria for programs. The principal author of this article, Dr. Lassar Gotkin, directs the Center's field testing and implementation studies.)

Educators interested in the field of programed instruction are writing letters to the Center asking for guidance in choosing programs. The answers to their questions must deal with separate but related problems: 1) the evaluation of programs; 2) the kinds of research educators can hope to carry out successfully in their schools. This article deals with the first; an article in the next issue will consider the second.

1. Is the Content Appropriate?

The title of the program does not necessarily supply this information. A training director of a communications company found that his students, on the average, completed a Basic Electricity program in half the time required for conventional instruction. On the test at the end, however, students using the programed text scored 30% below the average score of those taught conventionally. Some of the topics covered in the course, and thus in the text, were not covered in the program. Even where a publisher provides a thorough breakdown of the program's content, the educator must examine the material himself.

At present some programs are being selected for reasons other than the acceptability of their content. One elementary school principal believed enough in the possibilities of programed instruction and was so interested in learning about the field that he adopted a program about whose content he had reservations. As competing programs covering similar content become available, educators will seek to make their selections on the basis of field-test evidence. The remaining questions considered in this article will, therefore, pertain to data evolving from the field trials of programs with student groups.

2. Is the Content Well Programed?

The technical quality of a program is usually measured by analysis of students' responses reduced to a numerical "error rate." For the programer refining and revising his product, the analysis of students' responses is essential. The programer gains clues as to his selection of step size, of prompts, of ways to introduce new responses; generally he discovers whether or not he has communicated to the learners.

But the "error rate" itself is insufficient evidence of quality; it is perfectly possible to have low error rates and teach little. In the field trial of *English 2600,* using Denver 10th graders, the lowest of three ability groups showed an error rate of only 6.5%—but their performance on a test covering the material improved, on the average, by only two points. Teachers in conventional classes were equally unsuccessful with a similar group, but an error-rate analysis alone would have been most misleading to anyone debating the introduction of this program for low-ability students.

Error rates are affected by the programers' definition of the acceptable responses. A 5th grade program on the use of the microscope showed a 70% error rate on part 3 of the following frame:

Frame 6. This is a picture of one [ ]
Notice that it is thicker in the [ ] and thinner at the two [s]

The learners had no trouble filling in [lens] and [middle,] but 70% wrote [ends] instead of [edges]. The students' failure to give the response the programer desired is probably related to the two-dimensional presentation of the lens. It would seem, however, that the students grasped the concept. If the programer chose to accept ends as a correct response, though lenses do not have ends, the error rate on the item would be 0%.

An educator can also be misled by the predominant practice of reporting error rates in gross percentages across entire programs. A 10% error rate can result from very different distributions of errors. Finding that the test population errs at the 10% level *on all items* is far different from finding no errors on most items but large proportions of errors on a few, especially if these few are critical frames.

Differential error rates may be an appropriate criterion in determining whether or not a program is suitable for a given group of students. If a class of 8th graders has a 10% or less error rate on 80% of the items, while a comparable group of 7th graders has a 10% or less error rate on only 40% of the items, the educator should question the suitability of the program for the younger group as a whole.

Unless considered with other data, error analysis is valuable mostly as an indicator to the consumer that the program has at least been tested.

3. What Have the Students Learned From The Program?

Far more important than error analysis are the posttest results indicating what the students know upon completion of the program. And posttest results alone are not enough. With some programs, such as Eigen's *Sets, Relations, and Functions,* pretest information is usually unnecessary because the material is wholly new to the student. However, in the less exotic areas of study, students may have considerable knowledge of the subject before beginning a program. In a 4th-grade experiment using 100 words of a spelling program, 28% of the students knew at least 90 words before starting. The field test of a vocabulary building program with 7th graders showed considerable evidence not only of previous knowledge, but of previous knowledge that was incorrect. Post-

Continued on page 6

Faulty Frames

"Faulty Frames" began in the last issue of this Bulletin with a faulty "frame"! The sequence on the prefix *anti,* as given in the Bulletin, was not the complete sequence presented to students, nor was it an early sequence in the program. I had hoped to indicate this by numbering the items in the illustrations as 29, 31, 32, 34, 35. This very weak "prompt" failed.

(For those readers who wrote in and were concerned about what students would have been led to do had the sequence been the whole sequence, I should say that in the complete sequence neither the response "Antiarctic" nor the response "sympathy" occurred among students with adequate reading skills. Students had previously been taught when anti- *is changed to* ant- *and when prefixes ending in* n *are spelled with an* m. *The behavior required of the student at this late point in the program was complex and dependent on lengthy previous instruction which the student had in fact received, a fact I neglected to mention in my text.)*

Why were the items in the illustrative sequence *not* consecutive in the real program? If a certain number of items is required to get to a point, why not put the items in a straight line and get to the point? One of the commonest errors of new programers is to do just that.

An incomprehensible item will be spotted immediately, if not by the response of your first student, certainly by the responses of your first group of students. Unclear presentations lead to random errors and a bunching of omissions of responses. A faulty analysis of the subject matter, as illustrated in the last issue, will appear in a piling up of a consistent and unusual error, pointing a finger at the programer's mistake. But a sequence of six or eight items teaching a single new response will elicit few or no errors. Why, then, since we desire correct responding, is such a sequence called a programing mistake?

For the sake of convenience, the following example of the faulty sequence will use the same subject matter as the example in the last issue.

1. An anti-Communist is a person who works against the ideas of the Communists. The prefix meaning "against" is _____.
2. A person who was "against" the Nazi party in Germany was an _____-Nazi.
3. A kind of germ-killer that works against germs is an _____ septic.
4. A remedy that works against certain types of poisons is an _____dote.
5. A person who fights against society by misbehaving is said to be _____social.
6. A strong feeling against someone (the opposite of sympathy) is _____pathy.

Students will make almost no errors on this sequence. The rote responding involved here will prevent even the "intelligent error" *antisympathy*. Second, students will scan the later items rather than read them carefully. Third (a prediction verified by data), when the programer does change the subject, requiring a new response, a large number of students will write "anti" in the item no matter how absurd this response may be in the new context. Students have stopped reading. Fourth, if item 6 is repeated as a review or test at a later date, a large number of students will fail to answer it. They have not learned, although they have responded correctly. They have been responding correctly for THE WRONG REASON.

When errors occur, not in the sequence itself, but in review items or on the posttest, reasonable people might conclude that the students have not been given enough practice on the particular response. Programers confronted with such data tend to lengthen the sequence, providing more practice, which compounds the mistake. Even when such a repair does lead to competent behavior, it is at the cost of efficiency. A more appropriate response to such data is to break up the sequence. We want to *schedule,* not multiply, the student's practice.

A sequence on any subject proceeds from "simple" to complex items. The words the students is constructing are increasingly less familiar, the last one in item 6 being unfamiliar to most 8th graders, for whom the program was written. It would appear that the programer is withdrawing the prompt (or cue) provided by familiarity with the resulting word. In the last item, the student should respond "anti" because he now knows its definition. Item 6 appears to be unprompted, depending completely on the student's knowledge of the subject. On the contrary, it is *more* strongly prompted than item 2. Sequence itself is a prompt, and a powerful one.

The sequence prompt, though its name may be new, is not new to teachers. Whether we call the phenomenon "massed practice" as psychologists do, or "one form of cramming" as students do, its results are the same—momentary brilliance followed by rapid disintegration. Like any form of instruction, programed instruction must take account of the necessary scheduling of practice if the resultant behavior is to be permanently learned.

The repair of such a sequence depends, of course, on the specific subject matter. For the particular subject illustrated, the following pattern of insertions allows the programer to schedule the *"anti"* items. (Giving specific insertion items would require too much space.)

1. An anti-Communist is a person who works against the ideas of the Communists. The prefix meaning "against" is _____.
2. A person who was "against" the Nazi party in Germany was an _____-Nazi.

(Insert sequence: prefix *pro-* meaning "for," reviewing *pro-* meaning "forward" as in 'proceed'; contrast *pro-* and *anti-* words.)

10. A kind of germ-killer that works against germs is an _____ septic.

(Insert sequence: when is *anti* followed by a hyphen, when not.)

15. A remedy that works against certain types of poisons is an _____dote.

(Insert prefix *a-*, contrast with *anti-*.)

23. A person who fights against society by misbehaving is said to be _____social.

(Insert sequence: prefix *syn-* spelling changes.)

33. A strong feeling against someone (the opposite of sympathy) is _____pathy.

Continued on page 6

FAULTY FRAMES (Continued from page 5)

The resulting lesson, including work on the roots of the words, may reach 35 or more items. The same 35 items, of course, could be rearranged to take up each point by itself. By interlocking the separate developments, the programer requires the student to read each item, and to retain the information for increasingly lengthy periods of time. Many subject matters lend themselves to this type of treatment.

The advice to programers, then, is to break up sequences which require the same or similar responses. Lace such sequences with parallel developments of other parts of the material. If your review items or posttest data lead you to believe that you need to give students more and more practice, take a good look at the way you have sequenced your items. It may be true that students need more practice, but that they need only to have their practice better scheduled. If each step forward requires a little stretching of the "memory muscle," permanent retention will be more efficiently attained.

SUSAN MEYER MARKLE

Director of Programing, Center for Programed Instruction. Author of Programs in English Vocabulary, Critical Reading, and Programed Primer on Programing.

P. I.—E. or E. (Continued from page 3)

zations. Valid general conclusions can seldom be reached through such comparisons because the results depend on how good the program is and how good the teacher of the conventional method." This type of polemic "research" simply leads to further confusion and competitive feelings rather than to clarification and a climate of cooperation so necessary to scientific progress.

In conclusion, then, if we are ever going to be able to answer questions about the ultimate place of programed instruction in education by pointing to its positive and wide-ranging effects on our schools, we had better be sure that the scope of our method is wide enough to make such an effect possible.

Ideally programed instruction can help the schools of this country do better those things which they decide are the better things to do. To accomplish this, programed instruction must be both flexible and expandable. However, if one-sidedness becomes characteristic of programed instruction, education will justifiably judge programed instruction to be both inflexible and expendable.

There is no *single* way to program all subject matter for all students just as there is no one best way to teach all students at all times. Perhaps the best teacher or programer is the one who knows what way of instruction is the better way for some students some of the time. Programed instruction must learn that there are no more absolutes in modern education than there are in modern science. For, as modern science is confronted with an expanding and evolving intelligible universe, modern education must deal with an expanding and evolving universe of intelligibility. Programed instruction must likewise evolve and expand if it is to find a place in the educational mainstream.

P. KENNETH KOMOSKI

President, Center for Programed Instruction, Teacher, Ancient History, Collegiate School, New York City.

CHOOSING A PROGRAM: HOW TO START
(Continued from page 4)

test achievement without reference to pretest abilities may present a distorted picture.

Tests used for such evaluation should be made available to educators purchasing a program. Silberman has pointed out that most criterion tests used for research in programed instruction have been short, hastily constructed tests. Evans, Glaser, and Homme obtained contradictory results for retention using different criterion tests. Test results for a certain chemistry unit indicated almost no learning until analysis of the program results showed that students missed test items requiring application of principles to contexts that students did not understand.

4. What are the Characteristics of the Student Population(s) Involved in Trial Runs?

It is not uncommon to find the effectiveness of a particular program proclaimed on the basis of startling results with one student whose intelligence and previous experience with the materials are not revealed. Such selling procedure is clearly reprehensible, but we are not much better off if we know only that the program was tested with a group of students of unspecified ability levels. Many programs are being tested in suburban schools with students whose mean I. Q. is 115. Since posttests have repeatedly produced significant correlations with I. Q. and other measures of ability, very different results may be expected when a program is used with different ability levels.

Actually, the problems of identifying the appropriate intelligence level(s) for programs seem to be more complicated than those involved in finding the appropriate grade level(s). In the Center's "Research Summary No. 1," no differences in learning are reported for 9th, 10th, and 11th graders of above-average ability—high-ability students learned from the program regardless of grade level. Where *English 2600* was used on a single grade level, the highest ability group, starting from a high pretest score, learned the most; the lowest ability level, which had the most to learn, gained little. Educators may find it more appropriate to know the student's I. Q. and reading level rather than his grade level in deciding whether or not the student can handle a program.

To learn more about this issue, we plan to test a program requiring 8th grade reading comprehension with 4th through 11th graders. The 4th graders using the program are academically talented and on the average read at the 8th grade level while most of the 11th graders are below grade level in reading comprehension.

Some publishers seem ready to provide, with their programs, data similar to the normative data provided for standardized tests. Even when normative data are provided, however, educators might wish to check out the program for themselves; no school system is exactly like any other. Replication of field tests by educators will be discussed in the next article.

RECENT RESEARCH

Silverman, R. E. & Alter, M. "Response Mode, Pacing, and Motivational Effects in Teaching Machines." Department of Psychology, New York University. June, 1961.

The two experiments reported here were among five experiments published by the authors as one report.

I. This experiment compared the response and no-response conditions in programed text and machine formats.

Subject Matter:

A 90-frame program dealing with the binary numbers system. The "response" program was a standard Skinner-type fill-in-the-blank program. The "no response" program consisted of one or several complete sentences with no blanks.

Subjects:

60 college students required to participate in psychology experiments.

Procedure:

Subjects were screened on the basis of a preliminary test and then equal numbers of subjects were randomly assigned to machine and programed text groups; and to response or no-response conditions within each group. After completing the programs, subjects were given a final test consisting of constructed response questions.

Results:

No significant differences either between the response and no-response conditions or between the machine and text groups or interaction between the variables was found.

Comment:

Another study by these authors used a basic electricity program. Here too, there were no significant differences between the response and no-response conditions.

II. This experiment investigated the relative effectiveness and the motivational properties of three types of presentation devices: (a.) an elaborate electromechanical machine, (b.) a crude spool-type machine, and (c.) a programed text.

Subject Matter:

Two 90-frame Skinner-type programs: the binary number program cited above, and a portion of the Holland-Skinner program in psychology.

Subjects:

60 paid volunteers: 30 high school students and 30 college students.

Procedure:

Subjects were not familiar with the material. Ten high school and ten college students were randomly assigned to each of the three training devices. Each subject read through one of the programs and was then recalled, after an average interval of one week, to read through the other program. A constructed response test was administered at the end of each program.

Results:

Analysis of the final test scores indicate that:

(1) the three devices were equally effective, and, (2) the subjects performed as well in session 2 as they had in session 1 on each of the three devices. Acclimation to the devices did not seem to affect the subjects' performance significantly, (3) college students performed significantly better than high school students and, (4) none of the variables interacted significantly with any of the others.

Comment:

It would appear, again, that the type of device used to present the program is not important. As the authors point out, however, there is some question as to whether two exposures is sufficient to allow novelty effect to wear off. In evaluating the motivational properties of "teaching machines", it would probably have been more desirable to give the subjects more than two exposures.

General Comment:

The research might have presented more information; e.g., time required to complete program, IQ, program error rate, etc.

The programs used are brief (87 to 90 frames), and the sampling procedure virtually guarantees subjects of relatively high motivation. For highly motivated students in a novel experimental situation all the conditions may be equally effective even though prolonged exposure in classroom surroundings, with wide variations in motivation, might show an advantage for one or another means of presentation. For these reasons, the authors are conservative in making recommendations based on these findings. Criticisms of this nature are applicable to many, if not most, research studies involving programed instruction; the failure to find significant differences does not necessarily prove the absence of such differences.

Roe, V. K., Case, H. W., and Roe, A., "Scrambled vs. Ordered Sequence in Auto-Instructional Programs." Teaching Systems Research Project, Department of Engineering, UCLA, 1961.

This study investigated the importance of sequence in a program.

Program:

Mathematics. 71 frames were taken from the program on elementary probability developed at the UCLA Department of Engineering.

Subjects:

36 freshman psychology students classified into upper, middle, and lower ability groups.

Procedures:

Subjects were divided at random within each group, half taking the ordered sequence and half the scrambled sequence. The design, is 3 x 2, with six subjects in each cell.

Each item was presented on a 4 x 6 card, together with a multiple-choice statement. Students responded by removing a patch of rubber-cement below the answer they thought was correct. After completing the program, each student was given an attitude questionnaire and a criterion test. Time taken to complete the program and time taken to complete the test were recorded.

Results:

Students in the upper ability groups showed lower error rates on the program and higher test scores than students in the lower ability groups. Both differences were significant. Within each group, however, there was no significant difference—in time required for learning error rate, or test score—between those who worked

Continued on page 8

RECENT RESEARCH

(Continued from page 7)

on the ordered sequence and those who worked on the scrambled sequence.

Comment:

In a previous study, a student who had failed to read the instructions in his programed text had worked through the program by reading down each page rather than by moving along the horizontal divisions from page to page. This student, nevertheless, got a high score on the criterion test. The researchers may have felt like Hugh Walpole's princes of Serendip, who ventured forth on a journey and made marvelous discoveries of things they were not seeking.

The authors' conclude that college level students may not require such careful sequencing of auto-instructional items as has previously been supposed." The authors further state, "It would be interesting to examine the scrambled vs. ordered sequence question with students of various age levels with programs of different length, and in different subject areas, and with retention tests made at spaced intervals of time." Agreed.

———o———

Gagne, R. M. and Dick, W., "Learning Measures in a Self-Instructional Program in Solving Equations." Department of Psychology. Princeton University, 1960.

This study investigated three dimensions of "what was learned" from a beginning algebra problem.

Program:

Mathematics. The program "Solving Equations," assumes knowledge of 6th grade arithmetic. The experiments identified 21 equation-solving concepts in the program. The desired terminal behavior was the solution of simple linear equations. 8 booklets. 245 frames.

Subjects:

Fifty-two 7th grade students of "middle-ability" in mathematics.

Procedures:

Students completed one booklet a day for 8 days. 20 to 50 minutes were required to finish lessons. Two internal variables—error rate and time required to complete the program—were measured for each subject. Three tests of learning—(1) Verbalization of concepts, (2) Performance in using the concepts, and (3) solution of unfamiliar equations—were administered immediately following completion of the program. Six weeks later, verbalization and performance tests were given again. No delayed Transfer Test was administered.

Results:

(1) Internal variables: error rate was 6.7%. Mean learning time was 267 minutes with the range 190 to 380. The standard deviation was 37 minutes.

(2) External variables: Verbalization Test mean was 14.0 (of 21) on the immediate and 13.0 on the 6-week delayed t Test mean was 10.5 ıe immediate and 9.5 on .ĸ-delayed test. These results indicate considerable retention, though the loss in both instances is statistically significant. Transfer Test: mean performance was 2.1 out of a possible 50. This result indicates little ability to solve new equations.

(3) Relationships among measures:

(a) Correlations between Transfer and other criterion tests ran from .29 to .44.

(b) the range of correct responses on the Performance Test ran from 14% to 98%. The 8 test items showing the lowest level of achievement (each item was missed by more than 50% of the students) involved the more complex procedures "taught" in the program. Only 14% of the subjects developed the terminal behavior of solving simple equations.

Comments:

Particularly valuable is the authors' introductory discussion of internal and external variables as criteria for evaluating programs.

Results indicate that "what was learned" was retained but not transferred to solving "new" equations, though the new equations were "not considered particularly difficult by the experimenter."

Given the low achievement on the procedural concepts of the Performance Test (only 14% could solve simple equations), the low transfer scores are not surprising. The authors state, "Regardless of the internal criterion measures . . . (the error rate was 6.7%) . . . the transfer test scores make one reluctant to state that the learning program has truly taught "equation solving." This study demonstrates the limitations of error rate as a criterion for assessing the effectiveness of programs and the importance of using meaningful tests of terminal behavior.

Center to Publish Research Journal

Effective this fall, CPI will publish the JOURNAL OF PROGRAMED INSTRUCTION. The Journal will incorporate and further the work of AUTOMATED TEACHING BULLETIN, which was started two years ago by Rheem Califone Corporation. The new publication will attempt to keep pace with the rapid acceleration of experiment in programed instruction. The intent is to keep JPI primarily a research information exchange medium for all those who are involved in programing research, and others interested in **research** developments in this field. Theoretical articles on programing will also be included from time to time. Rheem Califone feels that it is appropriate that this service be carried on by a non-profit organization.

PROGRAMED INSTRUCTION, the Center's information bulletin will maintain its present combination of news and articles on programing, plus research abstracts.

The editor of will be Lincoln F. Hanson, an experimental psychologist of the CPI staff. The protem editorial board will include Drs. John Barlow, Bernard Everett, James Holland, David Klaus, Douglas Porter and Leonard Silvern. Articles for the Journal should be directed to Dr. Hanson, c/o The Center.

A free initial distribution of the JPI will be made late this fall to the Rheem mailing list and others requesting it promptly. Subsequent issue will be available only a subscription basis. Researchers are invited to submit reports to JPI: such reports will be published as soon as possible after acceptance.

The Center for Programed Instruction, Inc.
365 West End Avenue at 77th Street
New York 24, New York

NON-PROFIT ORG.

U. S. POSTAGE
PAID
New York, N.Y.
Permit No. 3850